

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 05-108084

(43)Date of publication of application : 30.04.1993

(51)Int.Cl.

G10L 3/00

(21)Application number : 03-298247

(71)Applicant : RICOH CO LTD

(22)Date of filing : 17.10.1991

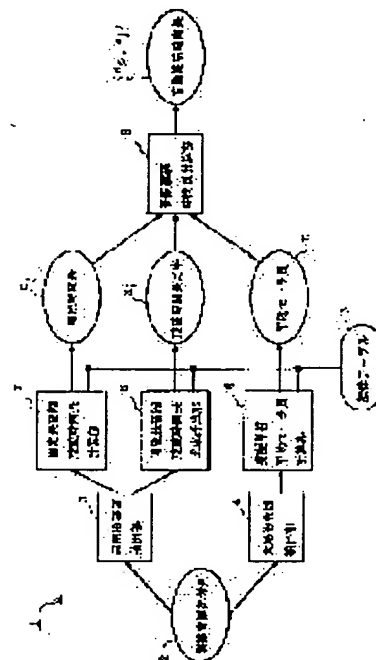
(72)Inventor : SAKAYORI TETSUYA

(54) SPEECH SYNTHESIZING DEVICE

(57)Abstract:

PURPOSE: To realize variation in speaking speed without spoiling the naturalness nor articulation of a phoneme and to synthesize a speech with more natural rhythm.

CONSTITUTION: A fixed length phoneme continuance calculation part 7 calculates the continuance of a fixed length phoneme. A variable length phoneme continuance ratio calculation part 6 calculates the continuance ratio of a variable length phoneme. A phoneme continuance calculation part 9 separately calculates the continuance of the fixed length phoneme and the continuance of the variable length phoneme for the continuance of speaking units according to the ratio of the continuance of the fixed length phoneme and the continuance of the variable length phoneme and performs ruled speech synthesis based upon the continuance values of the respective phonemes calculated for the continuance of the speaking units.



LEGAL STATUS

[Date of request for examination]

25.09.1998

[Date of sending the examiner's decision of rejection]

17.07.2001

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平5-108084

(43) 公開日 平成5年(1993)4月30日

(51) Int. Cl.⁵

G10L 3/00

識別記号

H 8946-5H

FI

審査請求 未請求 請求項の数4 (全6頁)

(21) 出願番号 特願平3-298247

(22) 出願日 平成3年(1991)10月17日

(71) 出願人 000006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72) 発明者 酒寄 哲也

東京都大田区中馬込1丁目3番6号 株式
会社リコー内

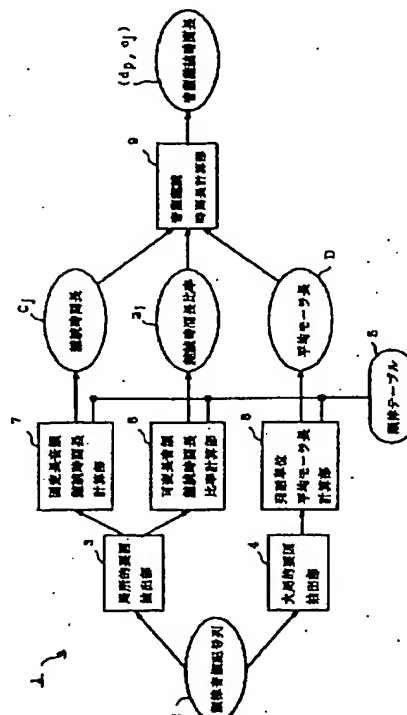
(74) 代理人 弁理士 植本 雅治

(54) 【発明の名称】 音声合成装置

(57) 【要約】

【目的】 自然な発話速度の緩急を音韻の自然性や明瞭性を損なうことなく実現することができ、一層自然なリズムの音声合成することの可能である。

【構成】 固定長音韻継続時間長計算部7では、固定長音韻の継続時間長を算出する。可変長音韻継続時間長比率計算部6では、可変長音韻の継続時間長比率を算出する。音韻継続時間長計算部9では、固定長音韻の継続時間長と可変長音韻の継続時間長比率とに基づき、発話単位の継続時間長に対する各音韻の継続時間長を固定長音韻と可変長音韻とに分けて算出し、発話単位の継続時間長に対して算出された各音韻の継続時間長により規則音声合成がなされる。



【特許請求の範囲】

【請求項 1】 固定長音韻の継続時間長に関する情報を算出する固定長音韻継続時間長算出手段と、可変長音韻の継続時間長に関する情報を算出する可変長音韻継続時間長算出手段と、固定長音韻の継続時間長に関する情報と可変長音韻の継続時間長に関する情報とに基づき、発話単位の継続時間長に対する各音韻の継続時間長を固定長音韻と可変長音韻とに分けて算出する音韻継続時間長算出手段とを有し、発話単位の継続時間長に対して算出された各音韻の継続時間長により規則音声合成がなされるようになっていて、これを特徴とする音声合成装置。

【請求項 2】 前記固定長音韻継続時間長算出手段は、固定長音韻の継続時間長に関する情報として、各固定長音韻の継続時間長を算出し、前記可変長音韻継続時間長算出手段は、可変長音韻の継続時間長に関する情報として、各可変長音韻の継続時間長比率を算出し、前記音韻継続時間長算出手段は、発話単位平均モーラ長に基づく発話単位全体の継続時間長から各固定長音韻の継続時間長の合計を差し引き、この差し引き結果を各可変長音韻の継続時間長比率によって分割して、それぞれの可変長音韻の継続時間長とするようになっていて、これを特徴とする請求項 1 記載の音声合成装置。

【請求項 3】 発話単位平均モーラ長、固定長音韻の継続時間長、可変長音韻の継続時間長比率のうちの少なくとも 1 つは、定性的あるいは定量的要因の線形結合によって算出されるようになっていて、これを特徴とする請求項 2 記載の音声合成装置。

【請求項 4】 発話単位平均モーラ長、固定長音韻の継続時間長、可変長音韻の継続時間長比率のうちの少なくとも 1 つは、定性的あるいは定量的要因を入力とする非線形結合ネットワークによって算出されるようになっていて、これを特徴とする請求項 2 記載の音声合成装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、音韻・韻律を表現する入力記号列に基づき規則音声合成を行なう音声合成装置に関する。

【0002】

【従来の技術】 従来、音韻・韻律を表現する入力記号列に基づき規則音声合成を行なう音声合成装置として、例えば特開平 3 - 1 8 8 9 9 号に開示されているようなものが知られている。この音声合成装置は、予め用意した音声素片のパラメータ系列を、音韻・韻律を表現する入力記号列に従って読み出し、音声パラメータ結合規則によって前記音声素片パラメータ系列を接続し、韻律制御規則によって音韻・韻律を表現する入力記号列に応じた音韻継続時間長、基本周波数パターン、振幅パターンなどを計算して、韻律を付加するようになっており、この際に、発声単位のモーラ数等の大局的要因を基に、発声単位全体の継続時間長を決定し、音韻環境等の局所的

要因を基に、発声単位内の各音韻の継続時間長比率を決定し、大局的要因によって決定した発声単位全体の継続時間長を、局所的要因によって決定された発声単位内の各音韻の継続時間長比率によって分割することにより、各音韻の継続時間長を決定するようになっている。このような音韻継続時間長制御がなされることにより、音韻継続時間長のバランスを保ち、自然な発話速度が実現でき、自然なリズムの音声を作成することができる。

【0003】

【発明が解決しようとする課題】 すなわち、上述した従来の音声合成装置では、大局的要因による発声単位全体の継続時間長と、局所的要因による音韻の継続時間長比率とを独立に扱うことによって、自然な発話テンポの緩急を表現することが可能となっている。しかしながら、発話単位内の各音韻は全て一様に伸縮されるので、音韻によっては自然性や明瞭性を損なうものがある。例えば、人間の発声を観察すると、母音や摩擦性の子音などは発話テンポの変化に伴って伸縮するのが普通である一方で、流音や半母音などは伸縮はほとんど起こらないが、上記の音声合成装置では、流音、半母音なども全て一様に伸縮されてしまうという問題があった。

【0004】 本発明はこのような問題を解決し、自然な発話速度の緩急を音韻の自然性や明瞭性を損なうことなく実現することができて、一層自然なリズムの音声を作成することの可能な音声合成装置を提供することを目的としている。

【0005】

【課題を解決するための手段】 上記目的を達成するために請求項 1 記載の発明は、固定長音韻の継続時間長に関する情報を算出する固定長音韻継続時間長算出手段と、可変長音韻の継続時間長に関する情報を算出する可変長音韻継続時間長算出手段と、固定長音韻の継続時間長に関する情報と可変長音韻の継続時間長に関する情報とに基づき、発話単位の継続時間長に対する各音韻の継続時間長を固定長音韻と可変長音韻とに分けて算出する音韻継続時間長算出手段とを有し、発話単位の継続時間長に対して算出された各音韻の継続時間長により規則音声合成がなされるようになっていて、これを特徴としている。

【0006】 また、請求項 2 記載の発明では、前記固定長音韻継続時間長算出手段が、固定長音韻の継続時間長に関する情報として、各固定長音韻の継続時間長を算出し、前記可変長音韻継続時間長算出手段が、可変長音韻の継続時間長に関する情報として、各可変長音韻の継続時間長比率を算出し、前記音韻継続時間長算出手段が、発話単位平均モーラ長に基づく発話単位全体の継続時間長から各固定長音韻の継続時間長の合計を差し引き、この差し引き結果を各可変長音韻の継続時間長比率によって分割して、それぞれの可変長音韻の継続時間長とするようになっていて、これを特徴としている。

【0007】 また、請求項 3 記載の発明では、発話単位

平均モーラ長、固定長音韻の継続時間長、可変長音韻の継続時間長比率のうちの少なくとも1つが、定性的あるいは定量的要因の線形結合によって算出されるようになっていることを特徴としている。

【0008】また、請求項4記載の発明では、発話単位平均モーラ長、固定長音韻の継続時間長、可変長音韻の継続時間長比率のうちの少なくとも1つが、定性的あるいは定量的要因を入力とする非線形結合ネットワークによって算出されるようになっていてることを特徴としている。

【0009】

【作用】請求項1、2記載の発明では、固定長音韻の継続時間長に関する情報と可変長音韻の継続時間長に関する情報とに基づき、発話単位の継続時間長に対する各音韻の継続時間長を固定長音韻と可変長音韻とに分けて算出するようにしており、これにより、自然な発話速度の緩急を音韻の自然性や明瞭性を損なうことなく実現することができる。

【0010】また、請求項3記載の発明では、発話単位平均モーラ長、固定長音韻の継続時間長、可変長音韻の継続時間長比率のうちの少なくとも1つを、定性的あるいは定量的要因の線形結合によって算出するようにしており、これにより、人間の発声に近い発話のリズムを簡単に得ることができる。

【0011】また、請求項4記載の発明では、発話単位平均モーラ長、固定長音韻の継続時間長、可変長音韻の継続時間長比率のうちの少なくとも1つを、定性的あるいは定量的要因を入力とする非線形結合ネットワークによって算出するようにしており、これにより、人間の発声に近い発話のリズムをより精密に得ることができる。

【0012】

【実施例】以下、本発明の一実施例を図面に基づいて説明する。図1は本発明に係る音声合成装置の一実施例のブロック図であり、図1には、音声合成装置の音韻継続時間制御部1の構成が示されている。

【0013】この音韻継続時間制御部1は、音韻、アクセント、声立て、ポーズなどを表わす記号からなる韻律音韻記号列2が入力するものとして構成されており、個々の音韻の継続時間長に影響を及ぼす要因、すなわち音韻環境、長音化促音化の有無、アクセント格の有無、発話単位内における位置等の局所的要因を韻律音韻記号列2から抽出する局所的要因抽出部3と、各音韻の継続時間長に直接影響を及ぼすものではなく発話単位全体の継続時間長に影響を及ぼすものと考えられる要因、すなわち発話単位のモーラ数、文法的・意味的重要性、システム使用者が要求する発話速度等の大局的要因を韻律音韻記号列2から抽出する大局的要因抽出部4と、韻律テーブル5と、局所的要因抽出部3で抽出された局所的要因に基づき可変長音韻の継続時間長比率 a_i を計算する可変長音韻継続時間長比率計算部6と、局所的要因抽出部

3で抽出された局所的要因に基づき固定長音韻の継続時間長 c_i を計算する固定長音韻継続時間長計算部7と、大局的要因抽出部4で抽出された大局的要因から発話単位の平均モーラ長 D を計算する発話単位平均モーラ長計算部8と、算出された可変長音韻の継続時間長比率 a_i 、固定長音韻の継続時間長 c_i 、発話単位の平均モーラ長 D に基づき全ての音韻の継続時間長を計算し制御する音韻継続時間長計算部9とを有している。

【0014】なお、ここで、発話単位とは、発話上一区切りとみなせる単位であり、音韻継続時間長を決定する上で一まとまりとして処理されるべき単位を意味している。また、可変長音韻の継続時間長比率とは、平均モーラ長に対する各可変長音韻の継続時間長の比であり、可変長音韻継続時間長比率計算部6は、例えば、韻律音韻記号から得られる各要因の線形結合によって可変長音韻の継続時間長比率を計算するようになっている。具体例としては、音韻環境などを要因アイテムとし、各可変長音韻継続時間長比率を外的基準として既知の数量化I類によって計算するようになっている。あるいは、可変長音韻継続時間長比率計算部6は、韻律音韻記号から得られる各要因を入力とする非線形結合ネットワークによって可変長音韻の継続時間長比率を計算することもできる。この具体例としては、音韻環境などを入力ユニットへの刺激として、可変長音韻継続時間長比率を出力ユニットからの反応とするようなニューラルネットワークを用いて、計算することができる。また、固定長音韻継続時間長計算部7、発話単位平均モーラ長計算部8も、可変長音韻継続時間長比率計算部6における上記のような計算手法とほぼ同様の手法を用いることによって、固定長音韻継続時間長、発話単位平均モーラ長を計算することができる。この際、図1の韻律テーブル5には、数量化I類を用いる場合の回帰係数や、ニューラルネットワークを用いる場合の重みなどが記憶されるようになっている。

【0015】また、音韻継続時間長計算部9は、発話単位の平均モーラ長 D に基づき、望ましい発話単位の継続時間長を算出し、また、各音韻の継続時間長については、発話単位の継続時間長とは独立に制御するようになっている。すなわち、各音韻の継続時間長については、例えば、発話単位の継続時間長が上記望ましい継続時間長となるように、これを計算し制御可能になっている。この際、各音韻の継続時間長の制御は、発話単位全体の発話テンポの変動に伴って継続時間長の変化する可変長音韻と発話テンポの変動によって継続時間長が変化しない固定長音韻とに分けて行なわれ、例えば、可変長音韻の継続時間長だけを伸縮し、固定長音韻の継続時間長は変化させないようにしている。

【0016】すなわち、本実施例では、発話単位平均モーラ長、可変長音韻継続時間長比率、固定長音韻継続時間長をそれぞれ別々に求めることにより、発話単位の継

続時間長と音韻の継続時間長とを独立に制御し、さらに、音韻の継続時間長を固定長音韻と可変長音韻とに分けて制御するようになっている。

【0017】次にこのような構成の音声合成装置、特に音韻継続時間制御部1の動作について説明する。韻律音韻記号列2が音韻継続時間制御部1に入力すると、音韻継続時間制御部1の局所的要因抽出部3、大局的要因抽出部4では、韻律音韻記号列2からそれぞれ局所的要因、大局的要因を抽出する。しかる後、発話単位平均モーラ長計算部8では、抽出された大局的要因に基づき上

述したような仕方では平均モーラ長Dを計算し、また、固定長音韻継続時間長計算部7、可変長音韻継続時間長比率計算部6は、発話単位内の各固定長音韻jの継続時間長 c_j 、各可変長音韻iの継続時間長比率 a_i を上述したような仕方

でそれぞれ計算する。
【0018】このようにして、発話単位の平均モーラ長D、発話単位内の各固定長音韻jの継続時間長 c_j 、各可変長音韻iの継続時間長比率 a_i がそれぞれ求まると、音韻継続時間長計算部9は、これらに基づいて全ての音韻継続時間長 d_i を算出する。

【0019】図2はこの様子を説明するための図であり、具体的には、発話単位内の可変長音韻、固定長音韻の個数をそれぞれn、mとすると、発話単位全体にわたる可変長音韻の継続時間の合計 S_1 、固定長音韻の継続時間長の合計 S_2 を、それぞれ次式によって求める。

【0020】

【数1】

$$S_1 = D \sum_{i=1}^n a_i$$

$$S_2 = D \sum_{j=1}^m c_j$$

【0021】ところで、発話単位のモーラ数をMとすると、望ましい発話単位継続時間長は、DMとなるが、上記可変長音韻の継続時間の合計 S_1 と固定長音韻の継続時間長の合計 S_2 との総合計値 $S (= S_1 + S_2)$ は、一般にDMとは一致しない。そこで、音韻継続時間長計算部9では、さらに総合計値SがDMと一致するように、各音韻の継続時間長を伸縮させる。但し、この際、可変長音韻の継続時間長についてだけ伸縮を行ない、固定長音韻の継続時間長 c_j については伸縮を行なわない。この結果、ある1つの可変長音韻の継続時間長 d_i は、次式によって求められる。

【0022】

【数2】

$$d_i = (DM - \sum_{j=1}^m c_j) a_i / \sum_{i=1}^n a_i$$

【0023】このように、各可変長音韻の継続時間長を数2により算出し、各固定長音韻の継続時間長については、 c_j をそのまま用いて、全ての音韻継続時間長を決

定することができる。このようにして得られた音韻継続時間長等により、韻律を付加して規則音声合成を行なうことができる。

【0024】この結果、得られる音声は、発話単位の継続時間長と音韻の継続時間長とが独立に制御されることによって、自然な発話テンポの緩急を実現することができ、さらに、各音韻を母音、子音の別とは無関係に、可変長音韻と固定長音韻とにわけて音韻継続時間長を制御することによって、音韻の自然性や明瞭性を損なうことなく自然な発話テンポの緩急を実現することができる。具体的には、母音や摩擦性の子音などの可変長音韻についてはこれを発話テンポの変化に伴なって伸縮させることができる一方、流音、半母音などの固定長音韻については、発話テンポが変化しても変化させずに済み、より自然なリズムの音声合成することができる。

【0025】また、発話単位の発話速度等や可変長音韻の継続時間長比率などを各種要因の線形結合によって計算することにより、人間の発声に近いリズムを簡単に得ることができる。また、これらを各種要因の非線形ネットワークによって計算することにより、人間の発声に近い発話のリズムをより精密に得ることができる。

【0026】

【発明の効果】以上に説明したように請求項1、2記載の発明によれば、固定長音韻の継続時間長に関する情報と可変長音韻の継続時間長に関する情報とに基づき、発話単位の継続時間長に対する各音韻の継続時間長を固定長音韻と可変長音韻とに分けて算出するようにしているので、自然な発話速度の緩急を音韻の自然性や明瞭性を損なうことなく実現することができ、一層自然なリズムの音声合成することができる。

【0027】また、請求項3記載の発明では、発話単位平均モーラ長、固定長音韻の継続時間長、可変長音韻の継続時間長比率のうちの少なくとも1つを、定性的あるいは定量的要因の線形結合によって算出するようにしているので、人間の発声に近い発話のリズムを簡単に得ることができる。

【0028】また、請求項4記載の発明では、発話単位平均モーラ長、固定長音韻の継続時間長、可変長音韻の継続時間長比率のうちの少なくとも1つを、定性的あるいは定量的要因を入力とする非線形結合ネットワークによって算出するようにしているので、人間の発声に近い発話のリズムをより精密に得ることができる。

【図面の簡単な説明】

【図1】本発明に係る音声合成装置の一実施例のブロック図である。

【図2】音韻の継続時間長の制御を説明するための図である。

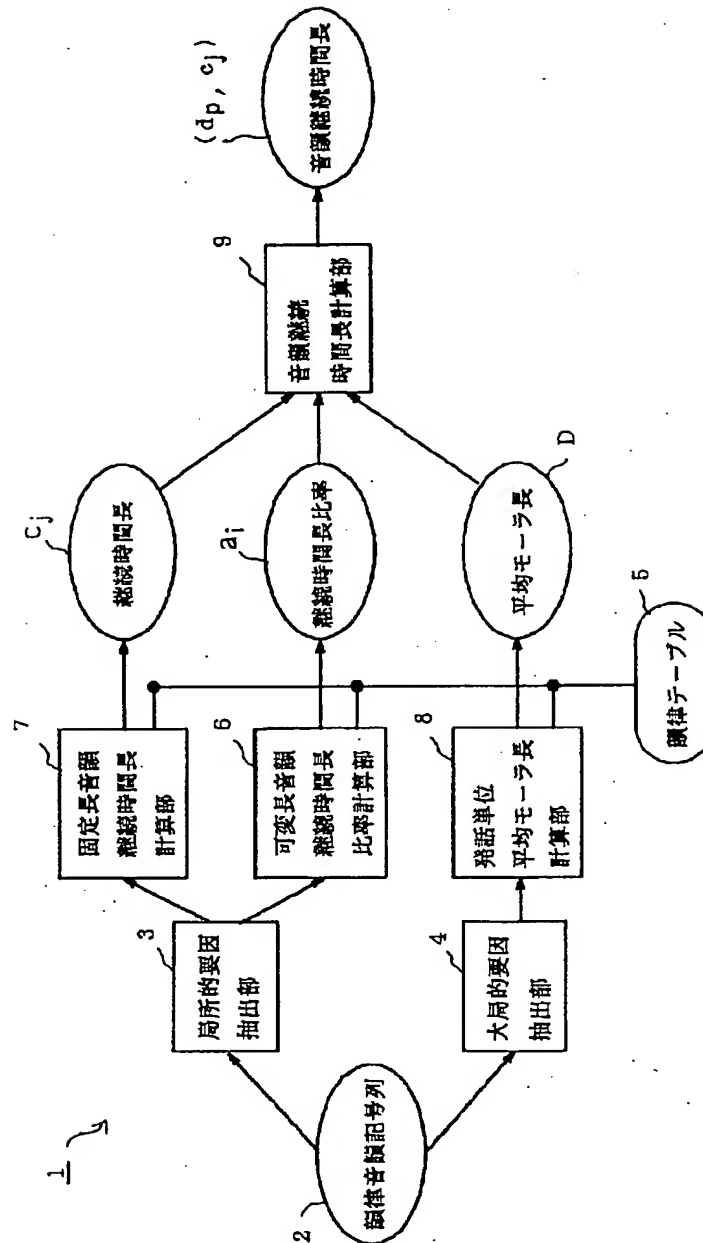
【符号の説明】

- 1 音韻継続時間制御部
- 2 韻律音韻記号列

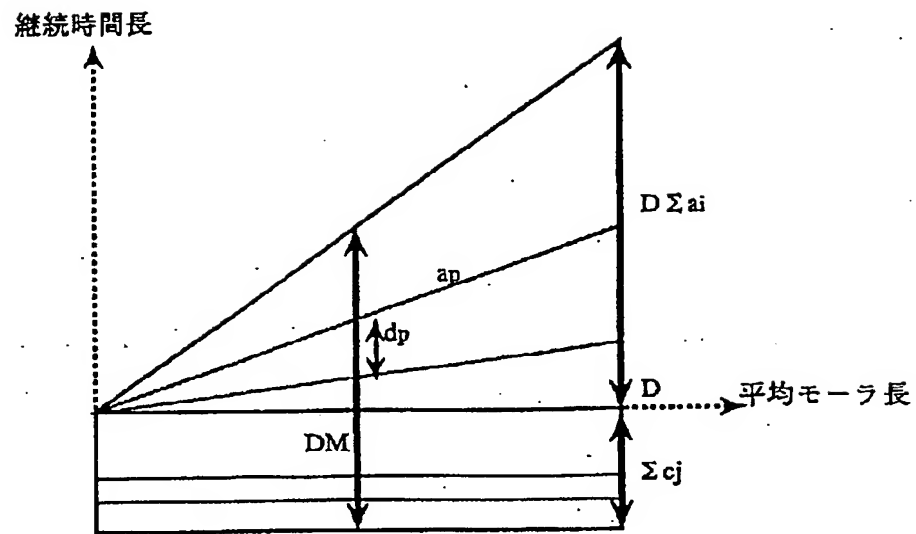
- 3 局所的要因抽出部
 4 大局的要因抽出部
 5 韻律テーブル
 6 可変長音韻継続時間長比率計算部

- 7 固定長音韻継続時間長計算部
 8 発話単位平均モーラ長計算部
 9 音韻継続時間長計算部

【図 1】



【図 2】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☒ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)